心理科学进展 2022, Vol. 30, No. 11, 2461–2472
Advances in Psychological Science
https://doi.org/10.3724/SP.J.1042.2022.02461

·研究方法(Research Method)·

纵向数据的调节效应分析*

方 杰[1] 温忠麟[2]

([1]广东财经大学新发展研究院/应用心理学系, 广州 510320)
([2]华南师范大学心理学院/心理应用研究中心, 广州 510631)

摘 要 目前调节效应检验主要是基于截面数据, 本文讨论纵向(追踪)数据的调节效应分析。如果自变量 X 和因变量 Y 有纵向数据, 调节效应可分为三类: 调节变量 Z 不随时间变化、Z 随时间变化、调节变量从自变量或因变量中产生。评介了基于多层模型、多层结构方程模型、交叉滞后模型和潜变量增长模型的纵向数据的多种调节效应分析方法。调节效应的分解和潜调节结构方程法的使用是纵向数据的调节效应分析的两大特点。对基于四类模型的调节效应分析方法进行综合比较后, 总结出一个纵向数据的调节效应分析流程。随后用实际例子演示如何进行纵向数据的调节效应分析, 并给出相应的 *Mplus* 程序。随后展望了纵向数据的调节效应分析的拓展方向, 例如基于动态结构方程模型的密集追踪数据的调节效应分析。

关键词 纵向数据, 调节效应, 多层模型, 多层结构方程模型, 交叉滞后模型, 潜变量增长模型
分类号 B841

调节(moderation)是社会科学研究中重要的方法学概念。如果自变量 X 对因变量 Y 的作用随第三个变量 Z 的变化而变化, 则称 Z 在 X 和 Y 之间起调节作用, 此时称 Z 为调节变量(见图 1)。近 30 年来, 调节研究备受关注, 分析方法和实际应用都得到长足发展。然而目前调节效应分析主要涉及截面数据(cross-sectional data), 许多研究者对截面数据的调节效应分析早已驾轻就熟, 但是, 当面对纵向(追踪)数据(longitudinal data)时, 如何进行调节效应分析呢? 近年来, 纵向数据的调节效应分析已经成为调节研究的新热点(方杰 等, 2018; 邱皓政, 2017; 温忠麟, 刘红云, 2020; Ozkok et al., in press; Preacher et al., 2016; Wen et al., 2014)。本文在评介多层模型(也称多水平模型, Multilevel Modeling, MLM)、多层结构方程模型(Multilevel Structural Equation Model, MSEM)、交叉滞后模型(Cross-Lagged Model, CLM)和潜变量增长模型(Latent Growth Model, LGM)的纵向数据的调节效应分析方法后, 对基于 4 类模型的调节效应分析方法进行了综合比较, 并提出一套纵向数据的调节效应分析流程; 然后用实际例子演示如何进行纵向数据的调节效应分析。最后展望了纵向数据的调节效应分析的拓展方向。

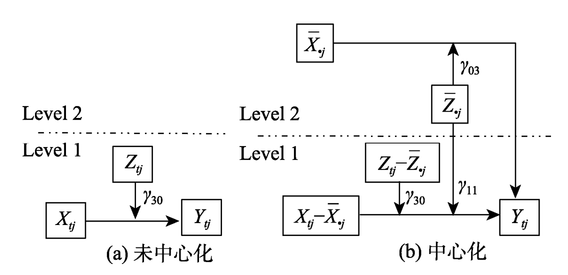


图 1 基于多层模型的纵向数据的调节模型

1 基于多层模型的纵向数据的调节效应分析

纵向(追踪)数据可看成是一个两层结构, 重复测量或测量点 t 为第 1 层, 观测个体 j 为第 2 层, 可使用多层模型进行纵向数据分析。如果调节变量不随时间变化(即时不变, 只需要一次测量),

收稿日期: 2021-10-16
* 国家自然科学基金项目(32171091)、国家社会科学基金项目(17BTJ035)资助。
通信作者: 温忠麟, E-mail: wenzl@scnu.edu.cn

则调节变量在层 2 (用 Z_j 表示), 自变量和因变量都在层 1, 可使用 2×(1→1)调节模型(这三个数字依次代表调节变量、自变量和因变量的层级, 数字 2 表示层级 2, 数字 1 表示层级 1)进行纵向数据的调节效应分析, 具体分析方法详见综述方杰等(2018), 不再赘述。但如果调节变量是随时间变化(即时变)的, 可以像自变量和因变量那样获取纵向数据(用 Z_{tj} 表示), 此时的调节效应分析复杂多了, 以下将详述相应的调节效应分析方法。

1.1 传统的多层调节效应分析

如果所有变量都随时间变化, 则所有变量都在层 1, 可使用 1×(1→1)调节模型进行纵向数据的调节效应分析。此时, 纵向数据的调节效应分析(图 1(a), Preacher et al., 2016)可表示为

$$层\text{-}1:\ Y_{tj} = \beta_{0j} + \beta_{1j}X_{tj} + \beta_{2j}Z_{tj} + \beta_{3j}X_{tj}Z_{tj} + \varepsilon_{tj} \quad (1)$$

$$层\text{-}2:\ \beta_{0j} = \gamma_{00} + \mu_{0j} \quad (2a)$$

$$\beta_{1j} = \gamma_{10} + \mu_{1j} \quad \beta_{2j} = \gamma_{20} + \mu_{2j}$$
$$\beta_{3j} = \gamma_{30} + \mu_{3j} \quad (2b)$$

将方程(2a)和(2b)代入方程(1)得:

$$Y_{tj} = \gamma_{00} + \gamma_{10}X_{tj} + \gamma_{20}Z_{tj} + \gamma_{30}X_{tj}Z_{tj} + \mu_{1j}X_{tj} + \mu_{2j}Z_{tj} + \mu_{3j}X_{tj}Z_{tj} + \varepsilon_{tj} + \mu_{0j} \quad (3)$$

调节项为 $X_{tj}Z_{tj}$ (见方程(3)), 调节效应的大小由回归系数 γ_{30} 表示, 如果回归系数 γ_{30} 显著不为 0, 则表示调节效应显著, 简单斜率为 $\gamma_{10}+\gamma_{30}Z_{tj}$。需要说明的是, 方程(2)中的 μ_{0j}、μ_{1j}、μ_{2j} 和 μ_{3j} 表示层 2 的残差, 说明截距 β_{0j} 和斜率 β_{1j}、β_{2j} 和 β_{3j} 在个体间可以变化(称为随机截距和随机斜率), 反映了个体间的差异。实际研究中, 为了简化模型, 研究者可将 β_{1j}、β_{2j} 和 β_{3j} 设为固定效应(即令 μ_{1j}、μ_{2j} 和 μ_{3j} 为 0)。

方程(1)~(3)所示的纵向数据的调节效应分析不仅操作方便, 而且容易解释, 方便应用。但这种方法有不足之处, 即层 1 变量 X_{tj} 将随时间变化的个体内(intra-individual)部分 $X_{tj}-\bar{X}_{\cdot j}$ 和不随时间变化的个体间(inter-individual)部分 $\bar{X}_{\cdot j}$ 混在一起, 同理 Z_{tj} 将随时间变化的个体内部分 $Z_{tj}-\bar{Z}_{\cdot j}$ 和不随时间变化的个体间部分 $\bar{Z}_{\cdot j}$ 混在一起(方杰等, 2018; Preacher et al., 2016)。

1.2 无混淆的多层调节效应分析

1.2.1 显变量的多层调节效应分析

为了解决前述混淆问题, Preacher 等(2016)建议用无混淆的多层模型(unconflated multilevel model)进行调节效应分析, 即将层 1 自变量 X_{tj}、Z_{tj} 按组均值 $\bar{X}_{\cdot j}$ 和 $\bar{Z}_{\cdot j}$ 中心化后进行建模, 同时将组均值 $\bar{X}_{\cdot j}$、$\bar{Z}_{\cdot j}$ 及其交互项 $\bar{X}_{\cdot j}\bar{Z}_{\cdot j}$ 置于层 2:

$$层\text{-}1:\ Y_{tj} = \beta_{0j} + \beta_{1j}(X_{tj}-\bar{X}_{\cdot j}) + \beta_{2j}(Z_{tj}-\bar{Z}_{\cdot j}) + \beta_{3j}(X_{tj}-\bar{X}_{\cdot j})(Z_{tj}-\bar{Z}_{\cdot j}) + \varepsilon_{tj} \quad (4)$$

$$层\text{-}2:\ \beta_{0j} = \gamma_{00} + \gamma_{01}\bar{X}_{\cdot j} + \gamma_{02}\bar{Z}_{\cdot j} + \gamma_{03}\bar{X}_{\cdot j}\bar{Z}_{\cdot j} + \mu_{0j} \quad (5a)$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}\bar{Z}_{\cdot j} + \mu_{1j} \quad (5b)$$

$$\beta_{2j} = \gamma_{20} + \mu_{2j} \quad \beta_{3j} = \gamma_{30} + \mu_{3j} \quad (5c)$$

将方程(5a)~(5c)代入方程(4)得:

$$Y_{tj} = \gamma_{00} + \gamma_{10}(X_{tj}-\bar{X}_{\cdot j}) + \gamma_{01}\bar{X}_{\cdot j} + \gamma_{20}(Z_{tj}-\bar{Z}_{\cdot j}) + \gamma_{02}\bar{Z}_{\cdot j} + \gamma_{30}(X_{tj}-\bar{X}_{\cdot j})(Z_{tj}-\bar{Z}_{\cdot j}) + \gamma_{03}\bar{X}_{\cdot j}\bar{Z}_{\cdot j} + \gamma_{11}(X_{tj}-\bar{X}_{\cdot j})\bar{Z}_{\cdot j} + \varepsilon_{tj} + \mu_{0j} + \mu_{1j}(X_{tj}-\bar{X}_{\cdot j}) + \mu_{2j}(Z_{tj}-\bar{Z}_{\cdot j}) + \mu_{3j}(X_{tj}-\bar{X}_{\cdot j})(Z_{tj}-\bar{Z}_{\cdot j}) \quad (6)$$

其中, 系数 γ_{10} 只表示变量 X_{tj} 随时间变化的个体内部分 $X_{tj}-\bar{X}_{\cdot j}$ 的效应, 系数 γ_{01} 只表示变量 X_{tj} 不随时间变化的个体间部分 $\bar{X}_{\cdot j}$ 的效应, 从而区分了变量 X_{tj} 的个体内和个体间效应(方程(6)与方程(3)对比)。同理, 系数 γ_{20} 只代表变量 Z_{tj} 随时间变化的个体内部分 $Z_{tj}-\bar{Z}_{\cdot j}$ 的效应, 系数 γ_{02} 只表示变量 Z_{tj} 不随时间变化的个体间部分 $\bar{Z}_{\cdot j}$ 的效应, 从而区分了变量 Z_{tj} 的个体内和个体间效应(方程(6)与方程(3)对比)。“无混淆”在本文中的含义是指, 可以区分变量的个体内和个体间效应。

纵向数据的调节效应也分为三个部分(图 1(b))。第一部分是随时间变化的个体内部分 $X_{tj}-\bar{X}_{\cdot j}$ 和 $Z_{tj}-\bar{Z}_{\cdot j}$ 的调节效应(以下称纯个体内×个体内(简写为 W×W)调节效应), 调节项为 $(X_{tj}-\bar{X}_{\cdot j})(Z_{tj}-\bar{Z}_{\cdot j})$, 调节效应的大小由回归系数 γ_{30} 表示, 如果回归系数 γ_{30} 显著不为 0, 则表示纯 W×W 调节效应显著, 简单斜率为 $\gamma_{10}+\gamma_{30}(Z_{tj}-\bar{Z}_{\cdot j})$。第二部分是不随时间变化的个体间部分 $\bar{X}_{\cdot j}$ 和 $\bar{Z}_{\cdot j}$ 的调节效应(以下称纯个体间×个体间(简写为 B×B)调节效应), 调节项为 $\bar{X}_{\cdot j}\bar{Z}_{\cdot j}$, 调节效应的大小由回归系数 γ_{03} 表示, 如果回归系数 γ_{03} 显著不为 0, 则表示纯 B×B 调节效应显著, 简单斜率为 $\gamma_{01}+\gamma_{03}\bar{Z}_{\cdot j}$。第三部分是随时间变化的个体内部分 $X_{tj}-\bar{X}_{\cdot j}$ 和不随时间变化的个体间部分 $\bar{Z}_{\cdot j}$ 的跨层次调节效应(以下称个体

间×个体内(简写为 B×W)调节效应), 调节项为 $(X_{tj}-\bar{X}_{\bullet j})\bar{Z}_{\bullet j}$, 调节效应的大小由回归系数 γ_{11} 表示, 如果回归系数 γ_{11} 显著不为 0, 则表示跨层次调节效应显著, 简单斜率为 $\gamma_{10}+\gamma_{11}\bar{Z}_{\bullet j}$。

需要说明的是, 方程 5(c)的 β_{2j} 和 β_{3j} 设为随机斜率便于研究者根据实际情况添加解释变量。例如, 将 $\beta_{2j}=\gamma_{20}+\mu_{2j}$ 改为 $\beta_{2j}=\gamma_{20}+\gamma_{21}\bar{X}_{\bullet j}+\mu_{2j}$, 则增加一个跨层次调节效应, 调节项为 $(Z_{tj}-\bar{Z}_{\bullet j})\bar{X}_{\bullet j}$, 如果回归系数 γ_{21} 显著不为 0, 则表示调节变量 $\bar{X}_{\bullet j}$ 对 $(Z_{tj}-\bar{Z}_{\bullet j})\rightarrow Y_{tj}$ 关系的跨层次调节效应显著。另外, 由于无混淆的多层模型将 X_{tj} 分为 $X_{tj}-\bar{X}_{\bullet j}$ 和 $\bar{X}_{\bullet j}$ 两个变量, 则可以将 $\bar{X}_{\bullet j}$ 当成调节变量来分析 $\bar{X}_{\bullet j}$ 对 $(X_{tj}-\bar{X}_{\bullet j})\rightarrow Y_{tj}$ 关系的调节效应, 此时并没有一个新的调节变量 Z (温忠麟, 刘红云, 2020)。

1.2.2 潜变量的多层调节效应分析

基于多层模型的纵向数据的调节效应检验中(方程(4)~(6)), 需要使用不随时间变化的个体间部分 $\bar{X}_{\bullet j}$ 和 $\bar{Z}_{\bullet j}$。多层模型设定变量为显变量, 即默认 t 次重复测量的 X_{tj} 和 Z_{tj} 的均值就等于 $\bar{X}_{\bullet j}$ 和 $\bar{Z}_{\bullet j}$, 但从 N 次重复测量中抽取部分重复测量结果($t\leqslant N$)来求 $\bar{X}_{\bullet j}$ 和 $\bar{Z}_{\bullet j}$, 会产生抽样误差(方杰, 邱皓政等, 2011; 方杰等, 2014, 2018)。比较好的解决方法是将 $\bar{X}_{\bullet j}$ 和 $\bar{Z}_{\bullet j}$ 看成是无法直接观测的潜变量, 使用多层结构方程模型进行分析, 能较好控制 $\bar{X}_{\bullet j}$ 和 $\bar{Z}_{\bullet j}$ 的抽样误差(方杰等, 2018; Preacher et al., 2016)。

多层结构方程模型的纵向数据的调节效应分析如图 2 所示。第一, 将每个层 1 测量的观测变量都分解为个体内部分和个体间部分两个潜变量(方杰, 邱皓政等, 2011; 方杰等, 2014, 2018; Preacher et al., 2016)。例如, 观测变量 X_{tj} 分解为个体内部分潜变量 R_{xtj} 和个体间部分潜变量 U_{xj}。同理, Z_{tj} 分解为两个潜变量 R_{ztj} 和 U_{zj}, Y_{tj} 分解为两个潜变量 R_{ytj} 和 U_{yj}。第二, 使用潜调节结构方程法(Latent Moderate Structural Equations, LMS, Klein & Moosbrugger, 2000), 根据

$$U_{yj}=\gamma_{00}+\gamma_{01}U_{xj}+\gamma_{02}U_{zj}+\gamma_{03}U_{xj}U_{zj}+\mu_{0j} \quad (7)$$

进行纯 B×B 的层 2 调节效应分析, 其中纯 B×B 调节项为 $U_{xj}U_{zj}$, 如果回归系数 γ_{03} 显著, 就表示纯 B×B 调节效应显著。同理, 根据

$$R_{ytj}=\beta_{0j}+\gamma_{10}R_{xtj}+\gamma_{20}R_{ztj}+\gamma_{30}R_{xtj}R_{ztj}+\varepsilon_{tj} \quad (8)$$

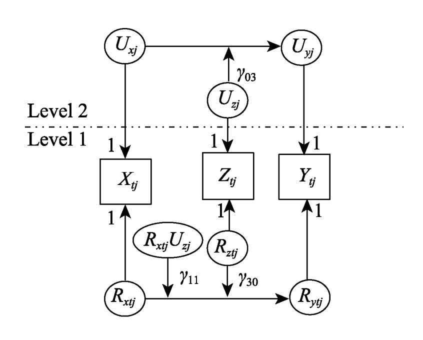


图 2 基于多层结构方程模型的纵向数据的调节模型

进行纯 W×W 的层 1 调节效应分析, 其中纯 W×W 调节项为 $R_{xtj}R_{ztj}$, 如果回归系数 γ_{30} 显著, 就表示纯 W×W 调节效应显著(见网络版附录程序)。第三, B×W 的跨层次调节效应分析可用两种方式进行分析。一种是随机系数预测法(Random Coefficient Prediction), 将方程(5b)变为 $\beta_{1j}=\gamma_{10}+\gamma_{11}U_{zj}+\mu_{1j}$, 做 β_{1j} 对 U_{zj} 的回归, 如果回归系数 γ_{11} 显著, 就表示跨层次调节效应显著(见网络版附录程序, 方杰等, 2018; Preacher et al., 2016)。另一种是潜调节结构方程法, 利用 R_{ytj}、R_{xtj}、U_{zj} 和调节项 $R_{xtj}U_{zj}$ 进行 B×W 的跨层次调节效应分析。就跨层次调节效应分析而言, 随机系数预测法的计算量比潜调节结构方程法要少(方杰等, 2018; Preacher et al., 2016), 所以建议选用随机系数预测法进行跨层次调节效应分析。

2 基于交叉滞后模型的纵向数据的调节效应分析

2.1 传统的交叉滞后的调节效应分析

传统的交叉滞后模型的调节效应分析(图 3)可表示为

$$X_{it} = \beta_X X_{i(t-1)}+\gamma_X Y_{i(t-1)}+C_X X_{i(t-1)}Y_{i(t-1)}+\varepsilon_{Xit} \quad (9)$$

$$Y_{it} = \beta_Y Y_{i(t-1)}+\gamma_Y X_{i(t-1)}+C_Y X_{i(t-1)}Y_{i(t-1)}+\varepsilon_{Yit} \quad (10)$$

其中, 下标 i 表示个体, 系数 β_X 和 β_Y 表示自回归效应(autoregressive effect, 即每个变量前后时间点的效应), 系数 γ_X 和 γ_Y 表示滞后效应(lag effect, 即每个变量对前因变量的回归)。调节项是 $X_{i(t-1)}Y_{i(t-1)}$, 如果系数 C_X 和(或) C_Y 显著不为 0, 则表示调节效应显著。值得注意的是, 自回归系数、滞后系数和调节系数的下标都没有时间 t, 表示自回归效应、滞后效应和调节效应在不同时间点都具有稳定性(目的是简化模型)。ε_{Xit} 和 ε_{Yit} 表示误差项, 服从正态分布。

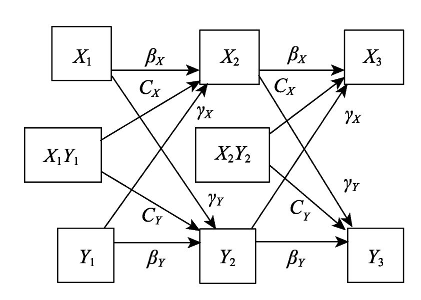


图 3 传统的交叉滞后调节模型

交叉滞后模型的调节效应至少有两个意义(以 C_X 为例进行说明)。第一, $X_{i(t-1)}$调节了滞后效应 $Y_{i(t-1)}\rightarrow X_{it}$。第二, $Y_{i(t-1)}$调节了自回归效应 $X_{i(t-1)}\rightarrow X_{it}$ (Ozkok et al., in press)。总之, 交叉滞后模型可以考察具有时间先后关系的变量间的调节效应, 这是交叉滞后模型的一大特色。值得说明的是, 在截面数据的调节研究中, 调节变量 Z 是一个不同于自变量 X 和因变量 Y 的变量。但是, 在纵向数据的调节研究中, 就不一定是这样了。例如, 在交叉滞后模型中, 自变量 X 和因变量 Y 变成了四个变量(X_{it}、$X_{i(t-1)}$、Y_{it} 和 $Y_{i(t-1)}$), 就可以不需要一个新的调节变量 Z 了(公式(9)和(10)), 这是纵向数据的调节研究有别于截面数据的一个特点。

2.2 无混淆的交叉滞后的调节效应分析

传统的交叉滞后模型的纵向数据的调节分析方法(见方程(9)和(10))同样存在不足, 即变量(如 X_{it} 和 Y_{it})也将随时间变化的个体内部分和不随时间变化的个体间部分混在一起。为了解决这一不足, Hamaker 等(2015)建议将中心化后的观测分数 X_{it} 分解为两个潜变量 X_{Bi} 和 X_{Wit} (图 4), 即

$$X_{it} = X_{Bi} + X_{Wit} \quad (11)$$

其中, X_{Bi} 表示不随时间变化(下标没有时间 t)的个体间差异, 即随机截距; X_{Wit} 表示随时间变化的个体内差异, 从而将变量 X_{it} 的个体内和个体间部分区分开。同理, 将中心化后的观测分数 Y_{it} 分解为两个潜变量 Y_{Bi} 和 Y_{Wit} (图 4), 即

$$Y_{it} = Y_{Bi} + Y_{Wit} \quad (12)$$

从而将变量 Y_{it} 的个体内和个体间部分区分开。此时, 方程(9)和(10)的调节项 $X_{i(t-1)}Y_{i(t-1)}$变为:

$$\begin{aligned} X_{i(t-1)}Y_{i(t-1)} &= (X_{Bi} + X_{Wi(t-1)})\times \\ &\quad (Y_{Bi} + Y_{Wi(t-1)}) = \\ &\quad X_{Bi}Y_{Bi} + X_{Bi}Y_{Wi(t-1)} + \\ &\quad Y_{Bi}X_{Wi(t-1)} + X_{Wi(t-1)}Y_{Wi(t-1)} \end{aligned} \quad (13)$$

可以看到 $X_{i(t-1)}Y_{i(t-1)}$实际上包含了 3 种调节效应(Ozkok et al., in press)。第一种是纯 W×W 调节效应, 即随时间变化的个体内调节效应 $X_{Wi(t-1)}Y_{Wi(t-1)}$; 第二种是纯 B×B 调节效应, 即不随时间变化的个体间调节效应 $X_{Bi}Y_{Bi}$, 第三种是 B×W 跨层次调节效应 $Y_{Bi}X_{Wi(t-1)}$和 $X_{Bi}Y_{Wi(t-1)}$。接下来将分别介绍这三种调节效应分析, 然后再介绍三种调节模型的整合分析。

2.2.1 纯个体内的潜调节效应分析

纯个体内的潜调节效应分析(图 4)可表示为方程(14)和(15)

$$\begin{aligned} X_{Wit} &= \beta_X X_{Wi(t-1)} + \gamma_X Y_{Wi(t-1)} + \\ &\quad C_X X_{Wi(t-1)} Y_{Wi(t-1)} + \varepsilon_{Xit} \end{aligned} \quad (14)$$

$$\begin{aligned} Y_{Wit} &= \beta_Y Y_{Wi(t-1)} + \gamma_Y X_{Wi(t-1)} + \\ &\quad C_Y X_{Wi(t-1)} Y_{Wi(t-1)} + \varepsilon_{Yit} \end{aligned} \quad (15)$$

其中, 调节项为 $X_{Wi(t-1)}Y_{Wi(t-1)}$, 如果系数 C_X 和(或) C_Y 显著, 表示调节效应显著(Ozkok et al., in press)。Ozkok 等(in press)用模拟研究证明, 纯个体内的潜调节效应分析(方程(14)和(15))优于传统的交叉滞后的调节效应分析(方程(9)和(10)), 表现为更小的偏差估计和更大的收敛率。Ozkok 等还用模拟研究发现, 纯个体内的潜调节效应分析优于纯个体内的显变量调节效应分析(即方程(11)中的 $X_{Wi(t-1)} = X_{i(t-1)}$−组均值 $\bar{X}_{\bullet i}$, $X_{Bi} = \bar{X}_{\bullet i}$; 同理, 方程(12)中的 $Y_{Wi(t-1)} = Y_{i(t-1)}- \bar{Y}_{\bullet i}$, $Y_{Bi} = \bar{Y}_{\bullet i}$)。

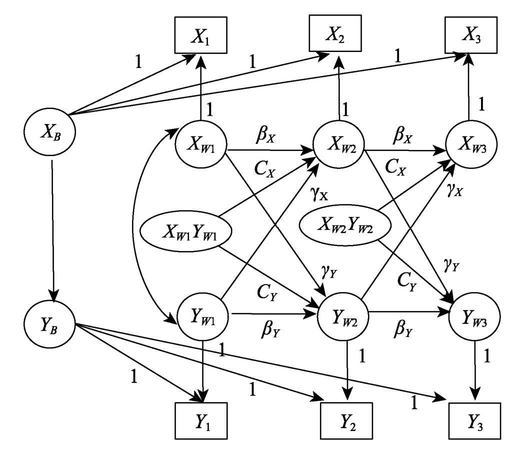


图 4 纯 W×W 的交叉滞后调节模型
(改编自 Ozkok et al., in press)

2.2.2 跨层次的潜调节效应分析

跨层次的潜调节效应分析(图 5)可表示为方程(16)和(17)

$$X_{Wit} = \beta_X X_{Wi(t-1)} + \gamma_X Y_{Wi(t-1)} + C_{1X} X_{Bi} Y_{Wi(t-1)} + C_{2X} X_{Bi} X_{Wi(t-1)} + C_{3X} Y_{Bi} X_{Wi(t-1)} + C_{4X} Y_{Bi} Y_{Wi(t-1)} + \varepsilon_{Xit} \quad (16)$$

$$Y_{Wit} = \beta_Y Y_{Wi(t-1)} + \gamma_Y X_{Wi(t-1)} + C_{1Y} X_{Bi} Y_{Wi(t-1)} + C_{2Y} X_{Bi} X_{Wi(t-1)} + C_{3Y} Y_{Bi} X_{Wi(t-1)} + C_{4Y} Y_{Bi} Y_{Wi(t-1)} + \varepsilon_{Yit} \quad (17)$$

其中，调节项为 $X_{Bi}Y_{Wi(t-1)}$、$X_{Bi}X_{Wi(t-1)}$、$Y_{Bi}X_{Wi(t-1)}$和 $Y_{Bi}Y_{Wi(t-1)}$，如果 C_{1X}、C_{2X}、C_{3X}、C_{4X}、C_{1Y}、C_{2Y}、C_{3Y}、C_{4Y} 中的某个系数显著，就表示与该系数相应的调节效应显著。与方程(14)和(15)相比，跨层次的潜调节效应模型(方程(16)和(17))仅仅是调节项变多了。Ozkok 等(in press)用模拟研究证明，跨层次的潜调节效应分析(方程(17)仅保留调节项 $Y_{Bi}X_{Wi(t-1)}$)也优于跨层次的显变量调节效应分析(即 $Y_{Bi} = \bar{Y}_{\cdot i}$、$X_{Wi(t-1)} = X_{i(t-1)} - \bar{X}_{\cdot i}$)，表现为 C_{3Y} 的偏差估计更小和更大的收敛率。

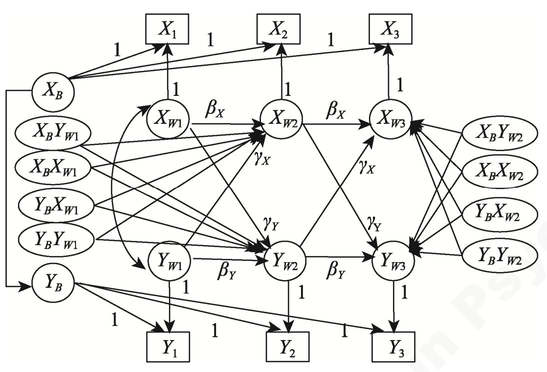


图 5　B×W 的交叉滞后调节模型
(改编自 Ozkok et al., in press)

2.2.3　纯个体间的潜调节效应分析

纯个体间的潜调节效应分析可表示为方程(18)和(19)

$$X_{Bi} = D_X Z_{Bi} + \gamma_X Y_{Bi} + C_X Z_{Bi} Y_{Bi} + \varepsilon_{Xi} \quad (18)$$

$$Y_{Bi} = D_Y Z_{Bi} + \gamma_Y X_{Bi} + C_Y Z_{Bi} X_{Bi} + \varepsilon_{Yi} \quad (19)$$

其中，Z_{Bi} 是一个新的个体间变量(如性别)，调节项是 $Z_{Bi}Y_{Bi}$ 和 $Z_{Bi}X_{Bi}$，如果 C_X 和(或)C_Y 显著，表示调节效应显著(Ozkok et al., in press)。

2.2.4　三种潜调节效应的整合分析

如果一个研究只有两个变量 X_{it} 和 Y_{it}，可以将两种潜调节效应整合在一起进行分析。具体地，由方程(14)和(15)进行个体内的潜调节效应分析，方程(16)和(17)进行跨层次的潜调节效应分析即可，无法进行个体间的潜调节效应分析。

如果自变量 X_{it} 和因变量 Y_{it} 随时间变化(纵向数据)，调节变量 Z_{Bi} 不随时间变化(截面数据)，就可以将三种潜调节效应整合在一起进行分析。具体地，由方程(14)和(15)进行个体内的潜调节效应分析，方程(18)~(20)进行个体间的潜调节效应分析(如果调节变量是二分类别变量，则去掉方程(20))，

$$Z_{Bi} = D_z Y_{Bi} + \gamma_Z X_{Bi} + C_z X_{Bi} Y_{Bi} + \varepsilon_{Zi} \quad (20)$$

跨层次的潜调节效应分析则需要在方程(16)和(17)中都各自加上两个调节项 $Z_{Bi}Y_{Wi(t-1)}$和 $Z_{Bi}X_{Wi(t-1)}$即可(Ozkok et al., in press)。Ozkok 等人用一个示例演示如何将潜调节结构方程法和 Bayesian 法相结合，对不随时间变化的调节变量(性别)做三种潜调节效应的整合分析。

如果所有变量都随时间变化(如 X_{it}、Y_{it} 和 Z_{it})，同样也可以将三种潜调节效应整合在一起进行分析。第一，方程(14)和(15)都各自增加 $Z_{Wi(t-1)}$、$X_{Wi(t-1)}Z_{Wi(t-1)}$和 $Y_{Wi(t-1)}Z_{Wi(t-1)}$三项，再加方程(21)，即可进行个体内的潜调节效应分析。

$$Z_{Wit} = \beta_Z Z_{Wi(t-1)} + \gamma_{ZX} X_{Wi(t-1)} + \gamma_{ZY} Y_{Wi(t-1)} + C_{XZ} X_{Wi(t-1)} Z_{Wi(t-1)} + C_{XY} X_{Wi(t-1)} Y_{Wi(t-1)} + C_{YZ} Y_{Wi(t-1)} Z_{Wi(t-1)} + \varepsilon_{Zit} \quad (21)$$

第二，由方程(18)~(20)进行个体间的潜调节效应分析。第三，方程(16)和(17)中都各自增加 $Z_{Wi(t-1)}$、$Z_{Bi}Y_{Wi(t-1)}$、$Z_{Bi}X_{Wi(t-1)}$、$Z_{Bi}Z_{Wi(t-1)}$、$Y_{Bi}Z_{Wi(t-1)}$、$X_{Bi}Z_{Wi(t-1)}$，再加上方程(22)，即可进行跨层次的潜调节效应分析。

$$Z_{Wit} = \beta_Z Z_{Wi(t-1)} + \gamma_{ZX} X_{Wi(t-1)} + \gamma_{ZY} Y_{Wi(t-1)} + C_{1Z} Y_{Bi} X_{Wi(t-1)} + C_{2Z} Y_{Bi} Y_{Wi(t-1)} + C_{3Z} Y_{Bi} Z_{Wi(t-1)} + C_{4Z} X_{Bi} Y_{Wi(t-1)} + C_{5Z} X_{Bi} X_{Wi(t-1)} + C_{6Z} X_{Bi} Z_{Wi(t-1)} + C_{7Z} Z_{Bi} Y_{Wi(t-1)} + C_{8Z} Z_{Bi} X_{Wi(t-1)} + C_{9Z} Z_{Bi} Z_{Wi(t-1)} + \varepsilon_{Zit} \quad (22)$$

需要说明的是，三种潜调节效应的整合分析会有很多个调节项，这无疑增加了数据不收敛的可能，也大大增加了运行时间。在实践中，建议研究者先从理论上排除一些没有充分理论依据的调节项，从而减少计算时间。

3　基于潜变量增长模型的纵向数据的调节效应分析

如果自变量、因变量和调节变量都随时间变化(纵向数据)，则潜变量增长模型的纵向数据的

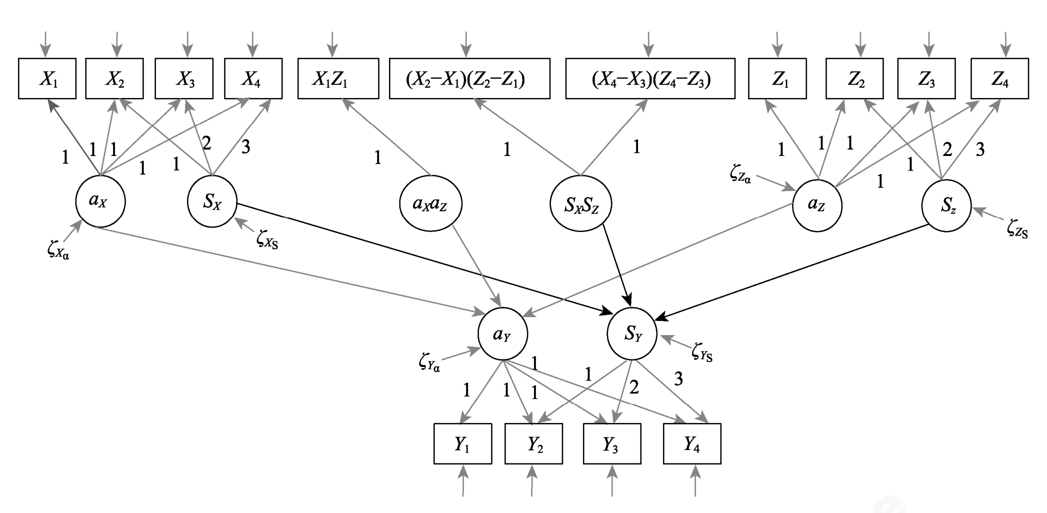


图 6 基于潜变量增长模型的纵向数据的调节模型(改编自 Wen et al., 2014)

调节效应分析包括四个潜变量增长模型，分别是自变量的潜变量增长模型、调节变量的潜变量增长模型、因变量的潜变量增长模型、和调节项的潜变量增长模型(图 6)。分析的关键是如何构建调节项的潜变量增长模型，Wen 等(2014)采用乘积指标构建了调节项的初始水平，用差积指标构建了调节项的变化率。具体地，自变量的潜变量线性增长模型可以表示为

$$X_1 = a_X + \delta_1 \quad (23a)$$
$$X_2 = a_X + S_X + \delta_2 \quad (23b)$$
$$X_3 = a_X + 2S_X + \delta_3 \quad (23c)$$
$$X_4 = a_X + 3S_X + \delta_4 \quad (23d)$$

其中，a_X 和 S_X 都是潜变量，各自表示变量 X 的初始水平(截距)和变化率。同理，调节变量的潜变量线性增长模型可以表示为

$$Z_1 = a_Z + \delta_5 \quad (24a)$$
$$Z_2 = a_Z + S_Z + \delta_6 \quad (24b)$$
$$Z_3 = a_Z + 2S_Z + \delta_7 \quad (24c)$$
$$Z_4 = a_Z + 3S_Z + \delta_8 \quad (24d)$$

将方程(23a)乘以(24a)，整理得：

$$X_1Z_1 = a_Xa_Z + a_Z\delta_1 + a_X\delta_5 + \delta_1\delta_5 \quad (25)$$

至此，初始水平的调节效应模型(图 6)，即 a_Z 对 a_X 和 a_Y 关系的调节效应可表示为

$$a_Y = b + da_X + ea_Z + fa_Xa_Z + \varepsilon_{aY} \quad (26)$$

其中，调节项为潜变量 a_Xa_Z，X_1Z_1 是 a_Xa_Z 的指标，由于是单指标潜变量，因此固定其负荷为 1。如果系数 f 显著，表示初始水平的调节效应显著(温忠麟，刘红云，2020; Wen et al., 2014)。

将方程(23b)减去(23a)、方程(24b)减去(24a)，得到：

$$(X_2 - X_1) = S_X + (\delta_2 - \delta_1) \quad (27a)$$
$$(Z_2 - Z_1) = S_Z + (\delta_6 - \delta_5) \quad (27b)$$

将方程(27a)乘以(27b)，整理得：

$$(X_2 - X_1)(Z_2 - Z_1) = S_XS_Z + S_X(\delta_6 - \delta_5) + S_Z(\delta_2 - \delta_1) + (\delta_2 - \delta_1)(\delta_6 - \delta_5) \quad (28)$$

同理，由方程(23d)、(23c)和方程(24d)、(24c)，得到：

$$(X_4 - X_3)(Z_4 - Z_3) = S_XS_Z + S_X(\delta_8 - \delta_7) + S_Z(\delta_4 - \delta_3) + (\delta_4 - \delta_3)(\delta_8 - \delta_7) \quad (29)$$

至此，变化率的调节效应模型(图 6)，即 S_Z 对 S_X 和 S_Y 关系的调节效应可表示为

$$S_Y = g + hS_X + kS_Z + wS_XS_Z + \varepsilon_{SY} \quad (30)$$

其中，调节项为潜变量 S_XS_Z，$(X_2 - X_1)(Z_2 - Z_1)$ 和 $(X_4 - X_3)(Z_4 - Z_3)$ 是 S_XS_Z 的指标(图 6)，如果系数 w 显著，表示变化率的调节效应显著(温忠麟，刘红云，2020; Wen et al., 2014)。

除了上述需要构建乘积项的乘积指标法，还可以使用无需构建乘积项的潜调节结构方程法进行基于潜变量增长模型的纵向数据的调节效应分析(详见第 6 部分的示例)。潜调节结构方程法的一个优势是，除了初始水平(调节项 a_Xa_Z)和变化率(调节项 S_XS_Z)的调节效应外，潜调节结构方程法还可以便利地产生其它调节项(如 a_XS_X、a_XS_Z 和 a_ZS_X)，灵活地进行其它调节效应分析。甚至当调节变量 Z 是不随时间变化时，潜调节结构方程法也可以便利地产生调节项(如 a_XZ、S_XZ)，进行调节

效应分析。另外, 由于潜变量增长模型将自变量 X 分为 a_X 和 S_X 两个潜变量, 从理论上说, 则可以将 S_X 当成调节变量, 利用潜调节结构方程法分析 S_X 对 $a_X \to S_Y$ 关系的调节效应, 并不一定需要一个新的调节变量 Z。

4　纵向数据的调节模型的综合比较

多层模型(MLM)、多层结构方程模型(MSEM)、交叉滞后模型(CLM)和潜变量增长模型(LGM)的调节模型比较如表 1 所示。纵向数据的调节效应分析的一大特点是调节效应的分解。具体地, 多层模型、多层结构方程模型和交叉滞后模型都将变量区分为随时间变化的个体内部分(W)和不随时间变化的个体间部分(B), 从而将纵向数据的调节效应分解为三部分, 即纯 W×W 调节、纯 B×B 调节和 B×W 跨层次调节。潜变量增长模型将纵向数据的调节效应分解为初始水平的调节效应和变化率的调节效应。

纵向数据的调节分析的另一个特点是进行潜变量的调节效应分析, 而且大都推荐使用潜调节结构方程(LMS)法进行潜变量的调节效应分析(Asparouhov & Muthén, 2021; Ozkok et al., in press; Preacher et al., 2016; Wen et al., 2014)。具体地, Preacher 等(2016)明确指出, 只有潜调节结构方程法能适用于 1 × (1→1)调节模型的 W×W 调节效应分析(即方程(8)的 $R_{xtj}R_{ztj}$)。Wen 等(2014)比较了乘积指标法和潜调节结构方程法在潜变量增长模型的变化率的调节效应检验中的表现。结果发现, 在正态条件下, 尤其是在小样本时, 潜调节结构方程法比乘积指标法在参数估计时产生更小的偏差。但是在数据非正态条件下, 乘积指标法比潜调节结构方程法在参数估计时产生更小的偏差。因此, 当数据正态分布条件下, 推荐使用潜调节结构方程法; 数据非正态条件下, 推荐使用乘积指标法。

目前, 潜调节结构方程法可通过 *Mplus* 软件方便实现(Muthén & Muthén, 1998-2017)。但是, *Mplus* 默认在潜调节结构方程法的调节效应估计中, 采用极大似然估计(Muthén & Muthén, 1998-2017)。当自变量和调节变量都是潜变量时, 极大似然估计必须用到数值积分(numerical integration), 这将导致计算量很大(Asparouhov & Muthén, 2021)。有研究者建议, 当自变量和调节变量都是潜变量时, 将潜调节结构方程法和 Bayesian 法相结合(*Mplus*8.3 及以后版本方可实现), 得到调节效应的 Bayesian 估计(Asparouhov & Muthén, 2021; Ozkok et al., in press)。Asparouhov 和 Muthén(2021)的模拟研究表明, 在潜调节结构方程法的调节效应分析中, 相比 Preacher 等(2016)使用的极大似然估计而言, Bayesian 估计会更快、更准确(表现为更小的绝对偏差(absolute bias)和更好的区间覆盖率(interval coverage rate))、且收敛率更高。

第三, 收集纵向数据进行分析的一个重要目的就是为了进行历时性因果推断(温忠麟, 2017), 如果要进行历时性因果推断, 在本文介绍的几类纵向模型中, 只有交叉滞后模型及其变体可以实现, 因为只有交叉滞后模型体现了历时性因果推测所必须的测量时间点先后关系, 同时考虑了自回归效应(方杰 等, 2021)。其他三种模型在建模过程中, 没有考虑变量之间影响的先后顺序, 其调节效应分析本质上与截面数据的调节分析无异, 要得出因果推论还要有统计以外的理据(温忠麟, 2017; 方杰等, 2021)。

另外, 如果是显变量的调节效应检验或想耗时少, 建议选择多层模型; 如果是测量次数只有两次, 建议选择交叉滞后模型; 如果是考察初始水平和(或)变化率的调节效应检验, 建议选择潜变量增长模型。

5　纵向数据的调节效应分析流程

面对一个纵向数据的调节效应分析任务, 研究者应当如何进行呢? 根据前面的讨论, 我们总

表 1　纵向数据的调节模型比较

模型	潜变量调节	调节效应的分解	测量次数	做历时性因果推断	用 LMS 法	耗时
MLM	×	W×W、B×B 和 B×W 调节	≥3	×	×	短
MSEM	√	W×W、B×B 和 B×W 调节	≥3	×	√	长
CLM	√	W×W、B×B 和 B×W 调节	≥2	√	√	长
LGM	√	初始水平调节、变化率调节等	≥3	×	√	长

注:“√”表示可以考察,“×”表示无法考察。

结出一套纵向数据的调节效应分析流程(见图 7)：

(1) 需要进行历时性因果推论吗？如果是，则用交叉滞后模型进行调节效应分析。否则，进入步骤 2。

(2)将纵向数据当成多层数据吗？如果是，则用多层结构方程模型进行调节效应分析。多层结构方程模型和多层模型更适合于描述个体间差异。否则，使用潜变量增长模型进行调节效应分析。只有潜变量增长模型可以同时研究某些变量对变化的影响(将变化作为因变量，如 S_Y)和变化如何影响其他变量(即把变化作为自变量，如 S_X) (温忠麟, 刘红云, 2020)。

(3)多层结构方程模型收敛吗？如果是，报告多层结构方程模型的调节效应分析结果。否则，使用多层模型进行调节效应分析。多层结构方程模型和多层模型的主要差别在于计算组均值时是否考虑抽样误差(见 1.2.2 部分)，但多层结构模型的收敛难度更大，因此优先使用考虑了抽样误差的多层结构方程模型，收敛不了再考虑多层模型。

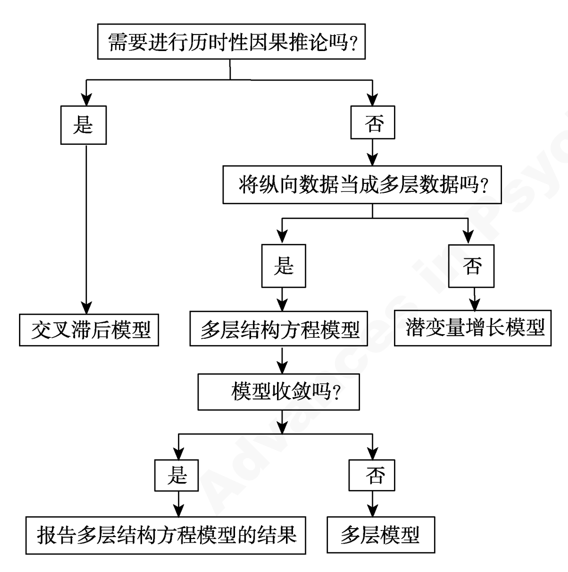


图 7 纵向数据的调节效应分析流程图

6 纵向数据的调节效应分析示例

下面用一个实际例子演示如何用图 7 的流程进行纵向数据的调节分析。本例的变量有家庭冲突(X)、青少年被欺负(Y)和烟酒使用，变量及其数据(每间隔半年重复测量 1 次，共重复测量 4 次，共 1132 名 11~15 岁的美国青少年)均来自 Little (2013)的研究。采用 *Mplus* 8.3 软件进行纵向数据的调节效应分析(*Mplus* 程序见网络版附录)。

6.1 基于交叉滞后的纵向数据的调节效应分析

如果需要进行历时性因果推论，则使用交叉滞后模型进行纵向数据的调节效应分析。为了简化模型，本例使用家庭冲突(X)和青少年被欺负(Y)两个变量，将 Bayesian 法和潜调节结构方程法相结合进行纯 W×W 的无混淆的交叉滞后模型的潜调节效应分析(方程(14)和(15)，有 3 个调节项 $X_{Wi1}Y_{Wi1}$、$X_{Wi2}Y_{Wi2}$ 和 $X_{Wi3}Y_{Wi3}$)。分析前，自变量和因变量按总均值进行了中心化。

第一，Bayesian 法是以马尔科夫链(Markov Chain)收敛为前提的，常用轨迹图(trace plot，如果多条马尔科夫链的轨迹高度重叠表明马尔科夫链收敛)和 PSR(Potential Scale Reduction，如果 PSR<1.1 (或更严格的标准 1.05)表明马尔科夫链收敛)来检验马尔科夫链是否收敛，更详细的介绍见综述方杰和张敏强等(2011)。轨迹图和 PSR 结果(TECH8 报告第 7700 次迭代时 PSR = 1.049)表明马尔科夫链收敛。

第二，纯 W×W 的潜调节效应分析结果显示，$X_{Wi(t-1)}\rightarrow X_{Wit}$ 的自回归效应 $\beta_X = 0.482$, $SE = 0.025$, 95%可靠区间[0.431, 0.532]不包含 0，自回归效应 β_X 显著，表明随着时间的推移，家庭冲突逐渐加重。$Y_{Wi(t-1)}\rightarrow X_{Wit}$ 的滞后效应 $\gamma_X = 0.003$, $SE = 0.02$, 95%可靠区间[−0.037, 0.044]包含 0，滞后效应 γ_X 不显著，表明 t−1 时刻的青少年被欺负不能显著预测后一时刻 t 的家庭冲突。$X_{Wi(t-1)}Y_{Wi(t-1)}\rightarrow X_{Wit}$ 的调节效应 $C_X = -0.053$, $SE = 0.026$, 95%可靠区间[−0.105, −0.004]不包含 0，调节效应 C_X 显著，表明 $X_{Wi(t-1)}$调节了滞后效应 $Y_{Wi(t-1)}\rightarrow X_{Wit}$，或者 $Y_{Wi(t-1)}$调节了自回归效应 $X_{Wi(t-1)}\rightarrow X_{Wit}$。

第三，$Y_{Wi(t-1)}\rightarrow Y_{Wit}$ 的自回归效应 $\beta_Y = 0.286$, $SE = 0.028$, 95%可靠区间[0.228, 0.341]不包含 0，自回归效应 β_Y 显著，表明随着时间的推移，青少年被欺负逐渐加重。这也符合实际，被欺负的青少年更容易继续被欺负。$X_{Wi(t-1)}\rightarrow Y_{Wit}$ 的滞后效应 $\gamma_Y = 0.07$, $SE = 0.028$, 95%可靠区间[0.016, 0.126]不包含 0，滞后效应 γ_Y 显著，表明 t−1 时刻的家庭冲突显著增加了后一时刻 t 的青少年被欺负。这可解释为，青少年和父母冲突越多，会让其他人认为，这个青少年缺少安全保护，会更容易将之作为欺负的对象。$X_{Wi(t-1)}Y_{Wi(t-1)}\rightarrow Y_{Wit}$ 的调节效应

C_Y = −0.039, SE = 0.033, 95%可靠区间[−0.105, 0.024]包含 0, 调节效应 C_Y 不显著。

第四, 不随时间变化的个体间关系 $X_{Bi}\to Y_{Bi}$ 的效应为 0.564, SE = 0.058, 95%可靠区间[0.459, 0.681]不包含 0, $X_{Bi}\to Y_{Bi}$ 的效应显著, 表明家庭冲突显著增加了青少年被欺负的风险。需要说明的是, 本例的纯 W × W 的潜调节效应检验(有 3 个调节项)耗时 3 个半小时, 因此, 跨层次的潜调节效应检验会耗时更长。

6.2 基于多层结构方程模型的纵向数据的调节效应分析

如果将纵向数据当成两层数据, 则使用多层结构方程模型进行调节效应分析。要研究的是烟酒使用(Z)对家庭冲突(X)和青少年被欺负(Y)的调节效应。采用 Bayesian 法和潜调节结构方程法、随机系数预测法相结合进行潜调节效应分析。其中, 随机系数预测法仅用来考察跨层次的潜调节效应, 其余潜调节效应用潜调节结构方程法进行考察。为了简化模型, 将 μ_{2j} 和 μ_{3j} 设为 0。

第一, 轨迹图和 PSR 结果(TECH8 报告第 900 次迭代时 PSR = 1.044)表明马尔科夫链收敛。第二, 基于多层结构方程模型的纵向数据的调节效应分析显示, 纯个体内(纯 W×W)的潜调节效应 γ_{30} = −0.191, SE = 0.077, 95%可靠区间[−0.35, −0.047]不包含 0, γ_{30} 显著不为 0, 表明纯个体内的潜调节效应显著。第三, 纯个体间(纯 B × B)的潜调节效应 γ_{03} = 0.232, SE = 0.086, 95%可靠区间[0.063, 0.394]不包含 0, γ_{03} 显著不为 0, 表明纯个体间的潜调节效应显著。第四, 跨层次的调节效应 γ_{11} = −0.113, SE = 0.156, 95%可靠区间[−0.403, 0.209]包含 0, γ_{11} 不显著, 表明跨层次的潜调节效应不显著。其它参数估计见表 2。

6.3 基于潜变量增长模型的纵向数据的调节效应分析

要研究的是烟酒使用(Z)对家庭冲突(X)和青少年被欺负(Y)的调节效应。采用 Bayesian 法和潜调节结构方程法相结合进行基于潜变量增长模型的潜调节效应分析(有二个调节项 a_Xa_Z 和 S_XS_Z)。

第一, 轨迹图和 PSR 结果(TECH8 报告第 1900 次迭代时 PSR = 1.011)表明马尔科夫链收敛。第二, 变化率的调节效应检验结果发现(以下呈现的是标准化解), S_X 对 S_Y 的效应显著(h = 0.363, SE = 0.084, 95%可靠区间[0.2, 0.528]不包含 0); S_Z 对 S_Y 的效应显著(k = −0.21, SE = 0.105, 95%可靠区间[−0.418, −0.007]不包含 0); 变化率的调节效应显著(w = 0.052, SE = 0.024, 95%可靠区间[0.006, 0.098]不包含 0), 即烟酒使用的变化率正向调节了家庭冲突的变化率与青少年被欺负的变化率的关系。

第三, 初始水平的调节效应检验结果发现(以下呈现的是标准化解), a_X 对 a_Y 的效应显著(d = 0.417, SE = 0.082, 95%可靠区间[0.264, 0.581]不包含 0); a_Z 对 a_Y 的效应不显著(e = −0.117, SE = 0.064, 95%可靠区间[−0.251, 0.001]包含 0); 初始水平的调节效应显著(f = −0.173, SE = 0.053, 95%可靠区间[−0.282, −0.074]不包含 0), 即烟酒使用的初始水平调节了家庭冲突的初始水平与青少年被欺负的初始水平的关系。

7 讨论和拓展

相比截面数据, 纵向数据的调节效应分析要复杂的多。对于自变量和因变量有纵向数据的调节效应分析, 可根据调节变量的不同类型分为三类。第一类是调节变量从自变量或因变量中产生。如多层模型中的 $\bar{X}_{\bullet j}$ 对 $(X_{tj}-\bar{X}_{\bullet j})\to Y_{tj}$ 关系的调节效应分析、交叉滞后模型中的纯个体内的潜调节效应分析(方程(14)和(15))、潜变量增长模型中的 S_X 对 $a_X\to S_Y$ 关系的调节效应分析。第二类是调节变量 Z 不随时间变化的调节效应分析。如多层模型中的 2×(1→1)调节效应分析、交叉滞后模型中的纯个体间的潜调节效应分析(方程(18)~(20))、潜变量增长模型中调节项为 a_XZ 或 S_XZ 的调节效应分析。第三类是调节变量 Z 随时间变化

表 2　多层结构方程模型的调节分析的部分参数估计结果

参数	γ_{00}	γ_{01}	γ_{02}	γ_{10}	γ_{20}	ε_{tj}	μ_{0j}	μ_{1j}
参数估计	1.545	0.445	0.009	0.196	−0.052	0.167	0.057	1.002
标准误	0.015	0.033	0.063	0.041	0.031	0.006	0.007	0.079
可靠区间下限	1.513	0.384	−0.128	0.120	−0.12	0.156	0.042	0.849
可靠区间上限	1.571	0.513	0.119	0.276	0.003	0.178	0.070	1.154

的调节效应分析。如多层模型中的 1×(1→1)调节效应分析、交叉滞后模型中的三种潜调节效应的整合分析、潜变量增长模型中调节项为 a_Xa_Z 或 S_XS_Z 的调节效应分析。

在阐述了基于多层模型、多层结构方程模型、交叉滞后模型和潜变量增长模型的多种纵向数据的调节效应分析后，对基于四类模型的调节效应分析方法进行了综合比较，接着总结出一个纵向数据的调节效应分析流程图，并用示例演示如何进行纵向数据的调节效应分析。但是，有关纵向数据的调节效应分析仍存在不足，尚需进一步深入研究和拓展。

第一，本文只考虑了自变量和因变量有纵向数据的情况。Montoya (2019)阐明了两水平被试内设计(仅因变量有纵向数据)的调节效应分析。具体地，研究者对同一组被试施加了两种不同水平的实验处理(自变量 X 是二分类别变量)，调节变量 Z 是不随时间变化的连续变量，因变量 Y 是随时间变化的连续变量(两种处理下的因变量 Y_1 和 Y_2)。两水平被试内设计的调节效应分析的方法是：由 $Y_{\text{diff}}=Y_2-Y_1$ 产生新变量 Y_{diff}，如果回归方程 $Y_{\text{diff}}=d_D+b_DZ+\varepsilon_D$ 的系数 b_D 显著，就表示调节效应显著。此外，如果自变量和调节变量都是不随时间变化的连续变量，因变量是随时间变化的连续变量，可以使用多层模型的 2×(2→1)调节效应分析。如果自变量是不随时间变化的连续变量，调节变量和因变量是随时间变化的连续变量，可以使用多层模型的 1×(2→1)调节效应分析。

第二，本文仅涉及测量次数较少($t\leqslant 10$)的纵向数据的调节效应分析，如果是重复测量次数较多($t\geqslant 10$)的纵向数据，常称为密集追踪数据(Intensive Longitudinal Data, 唐文清等, 2020; 郑舒方等, 2021)，如何进行纵向数据的调节效应分析呢？密集追踪数据常用动态结构方程模型(Dynamic Structural Equation Models, DSEM)进行数据分析(郑舒方等, 2021; McNeish & Hamaker, 2020)。动态结构方程模型是将交叉滞后模型和多层模型相结合，其本质是在多层模型中添加自回归效应和滞后效应(方程(31)和(32))。基于动态结构方程模型的密集追踪数据的调节效应分析可表示为

层-1： $$Y_{tj}=\beta_{0j}+\beta_{1j}Y_{(t-1)j}+\beta_{2j}X_{(t-1)j}+\varepsilon_{tj} \quad (31)$$

层-2： $$\beta_{0j}=\gamma_{00}+\gamma_{01}Z_j+\mu_{0j} \quad (32a)$$

$$\beta_{1j}=\gamma_{10}+\gamma_{11}Z_j+\mu_{1j} \quad (32b)$$

$$\beta_{2j}=\gamma_{20}+\gamma_{21}Z_j+\mu_{2j} \quad (32c)$$

其中，系数 γ_{10} 表示自回归效应，γ_{20} 表示滞后效应，系数 γ_{11} 表示不随时间变化的层 2 调节变量 Z_j 对自回归效应 $Y_{(t-1)j}\rightarrow Y_{tj}$ 的调节效应，系数 γ_{21} 表示调节变量 Z_j 对滞后效应 $X_{(t-1)j}\rightarrow Y_{tj}$ 的调节效应，如果系数 γ_{11} 和(或)γ_{21} 显著不为 0，就表示调节效应显著。在动态结构方程模型中，层 1 变量按组均值中心化，*Mplus* 软件自动将层 1 变量分解为个体内部分和个体间部分两个潜变量，Z_j 按总均值进行中心化(McNeish & Hamaker, 2020)。基于动态结构方程模型的调节效应分析的不足在于，目前仅能分析不随时间变化的调节变量 Z_j 的调节效应(Ozkok et al., in press)。

第三，本文仅涉及纵向数据的调节效应分析方法，已有研究还介绍了纵向数据的中介效应分析方法(刘国芳 等, 2018; 方杰等, 2021)，那么纵向数据的中介和调节效应整合在一起，如何进行纵向数据的有调节的中介效应分析呢？目前，纵向数据的有调节的中介效应检验可采用多层模型(Bauer et al., 2006)和多层自回归模型(Zhang et al., 2018)进行分析。但这两种方法都没有将变量分为随时间变化的个体内部分和不随时间变化的个体间部分。那么，在纵向数据的有调节的中介效应分析中，如何实现随时间变化的个体内效应和不随时间变化的个体间效应的有效分离呢？还有待深入研究。

纵向数据在心理学和其他社科领域研究中经常遇到，但不仅纵向数据的调节分析的方法学研究滞后于中介分析方法的发展，而且纵向数据的调节分析的应用研究在国内也很少，本文试图为应用研究者提供相应的方法学支持。实际上，纵向数据的调节效应的分析方法仍处蓬勃发展期，还有诸多问题亟待解决。例如，基于交叉滞后模型的纵向数据的调节效应分析，目前仅能提供非标准化解，那么如何提供标准化解呢？又如，纵向数据的调节效应分析目前主要通过 *Mplus* 软件实现，这在一定程度上限制了应用研究者进行纵向数据的调节效应分析，如何才能在其他软件上进行纵向数据的调节效应分析呢？再如，虽然潜调节结构法方程法优势突出，但很多应用研究者放弃该法的一个重要原因就是该法耗时长，如何才能减少潜调节结构方程法的耗时呢？方法的进步给研究者提供了一个深入理解和应用纵向数据

进行调节效应分析的机会，相信随着纵向数据的调节分析方法研究的深入，会不断增加我们对纵向数据的调节效应问题的理解。

参考文献

方杰, 邱皓政, 张敏强. (2011). 基于多层结构方程模型的情境效应分析——兼与多层线性模型比较. *心理科学进展*, *19*(2), 284–292.

方杰, 温忠麟, 邱皓政. (2021). 纵向数据的中介效应分析. *心理科学*, *44*(4), 989–996.

方杰, 温忠麟, 吴艳. (2018). 基于结构方程模型的多层调节效应. *心理科学进展*, *26*(5), 781–788.

方杰, 温忠麟, 张敏强, 任皓. (2014). 基于结构方程模型的多层中介效应分析. *心理科学进展*, *22*(3), 530–539.

方杰, 张敏强, 李晓鹏. (2011). 中介效应的三类区间估计方法. *心理科学进展*, *19*(5), 765–774.

刘国芳, 程亚华, 辛自强. (2018). 作为因果关系的中介效应及其检验. *心理技术与应用*, *6*(11), 665–676.

邱皓政. (2017). *多层次模式与纵贯资料分析——Mplus8 解析应用*. 台湾: 五南图书出版股份有限公司.

唐文清, 张敏强, 方杰. (2020). 时变效应模型及在密集追踪数据分析中的应用. *心理科学*, *43*(2), 488–497.

温忠麟. (2017). 实证研究中的因果推理与分析. *心理科学*, *40*(1), 200–208.

温忠麟, 刘红云. (2020). *中介效应和调节效应: 方法及应用*. 北京: 教育科学出版社.

郑舒方, 张沥今, 乔欣宇, 潘俊豪. (2021). 密集追踪数据分析: 模型及其应用. *心理科学进展*, *29*(11), 1948–1969.

Asparouhov, T., & Muthén, B. (2021). Bayesian estimation of single and multilevel models with latent variable interactions. *Structural Equation Modeling: A Multidisciplinary Journal*, *28*(2), 314–328.

Bauer, D. J., Preacher, K. J., & Gil, K. M. (2006). Conceptualizing and testing random indirect effects and moderated mediation in multilevel models: New procedures and recommendations. *Psychological Methods*, *11*(2), 142−163.

Hamaker, E. L., Kuiper, R. M., & Grasman, R. P. P. P. (2015). A critique of the cross-lagged panel model. *Psychological Methods*, *20*(1), 102–116.

Klein, A. G., & Moosbrugger, H. (2000). Maximum likelihood estimation of latent interaction effects with the LMS method. *Psychometrika*, *65*(4), 457–474.

Little, T. D. (2013). *Longitudinal structural equation modeling*. NY: Guilford Press.

McNeish, D., & Hamaker, E. L. (2020). A primer on two-level dynamic structural equation models for intensive longitudinal data in Mplus. *Psychological Methods*, *25*(5), 610–635.

Montoya, A. K. (2019). Moderation analysis in two-instance repeated-measures designs: Probing methods and multiple moderator models. *Behavior Research Methods*, *51*(2), 61–82.

Muthén, L. K., & Muthén, B. O. (1998-2017). *Mplus user's guide* (8th ed.). Muthén & Muthén.

Ozkok, O., Vaulont, M. J., Zyphur, M. J., Zhang, Z., Preacher, K. J., Koval, P., & Zheng, Y. (in press). Interaction effects in cross-lagged panel models: SEM with latent interactions applied to work-family conflict, job satisfaction, and gender. *Organizational Research Methods*.

Preacher, K. J., Zhang, Z., & Zyphur, M. J. (2016). Multilevel structural equation models for assessing moderation within and across levels of analysis. *Psychological Methods*, *21*(2), 189–205.

Wen, Z., Marsh, H. W., Hau, K-T., Wu, Y., Liu, H., & Morin, A. J. S. (2014). Interaction effects in latent growth models: Evaluation of alternative estimation approaches. *Structural Equation Modeling: A Multidisciplinary Journal*, *21*(3), 361–374.

Zhang, Q., Wang, L. J., & Bergeman, C. S. (2018). Multilevel autoregressive mediation models: Specification, estimation, and applications. *Psychological Methods*, *23*(2), 278–297.

Moderation analysis for longitudinal data

FANG Jie[1], WEN Zhonglin[2]

(1 *Institute of New Development & Department of Applied Psychology, Guangdong University of Finance & Economics, Guangzhou* 510320, *China*) (2 *Center for Studies of Psychological Application & School of Psychology, South China Normal University, Guangzhou* 510631, *China*)

Abstract: At present, the analysis of moderating effect is mainly based on cross sectional data. This article discusses how to analyze the moderating effect with longitudinal data. If the independent variable X and the dependent variable Y are longitudinal data, longitudinal moderation model can be divided into three categories according to the type of moderator: time-invariant moderator, time-variant moderator, and moderator generated from X or Y. Four types of longitudinal moderation analysis approaches are summarized: 1) Multilevel model (MLM); 2) Multilevel structural equation model (MSEM); 3) Cross-lagged model (CLM); 4) Latent growth model (LGM). It is found that the decomposition of the moderating effect and the use of latent moderating structural equation (LMS) method are the two characteristics of the moderation analysis for longitudinal data. A procedure for longitudinal moderation analysis is proposed. An example is illustrated with the software *Mplus*. Directions for future research on moderation analysis of longitudinal data are discussed, such as the moderation analysis for intensive longitudinal data based on the dynamic structural equation model.

Keywords: longitudinal data, moderation effect, multilevel model, multilevel structural equation model, cross-lagged model, latent growth model

附录:

基于交叉滞后模型的纯 W×W 调节效应分析的 Mplus 程序:

```
DATA:           FILE = 1.dat;
VARIABLE:       NAMES = x1-x4 m1-m4 y1-y4;
                USEVARIABLES = x1-x4 y1-y4;
DEFINE:    CENTER x1-x4 y1-y4 (GRANDMEAN);   !按总均值中心化
ANALYSIS:    ESTIMATOR = BAYES;     !使用 Bayesian 法
             TYPE = RANDOM;
             PROCESSORS = 8;           !多个处理器, 加快运算速度
             CHAINS = 2;               !两条马尔科夫链
             THIN = 2;                 !抽取后验分布的间隔
             FBITERATIONS = 30000;   !马尔科夫链迭代 3 万次
MODEL:     xb by x1-x4@1;   !设置 x 的个体间部分, 即随机截距
           yb by y1-y4@1;   !设置 y 的个体间部分, 即随机截距
           xw1 by x1@1;   x1@0.01;   !设置 x 的个体内部分
           xw2 by x2@1;   x2@0.01;
           xw3 by x3@1;   x3@0.01;
           xw4 by x4@1;   x4@0.01;
           yw1 by y1@1;   y1@0.01;   !设置 y 的个体内部分
           yw2 by y2@1;   y2@0.01;
           yw3 by y3@1;   y3@0.01;
           yw4 by y4@1;   y4@0.01;
           xb yb with xw1-xw4@0 yw1-yw4@0; !随机截距和个体内部分相关为 0

          yb on xb;                   !个体间效应
          xw4 on xw3(beta_x);     !自回归效应
          xw3 on xw2(beta_x);
          xw2 on xw1(beta_x);
          yw4 on yw3(beta_y);
          yw3 on yw2(beta_y);
          yw2 on yw1(beta_y);
          xw4 on yw3(gama_x);   !滞后效应
          xw3 on yw2(gama_x);
          xw2 on yw1(gama_x);
          yw4 on xw3(gama_y);
          yw3 on xw2(gama_y);
          yw2 on xw1(gama_y);
          xw1-xw4 yw1-yw4 WITH xw1-xw4@0 yw1-yw4@0; !  限制所有相关为 0
          xw1 with yw1; xw2 with yw2;    !允许自由估计的相关
          xw3 with yw3; xw4 with yw4;
          xwyw1 | xw1 XWITH yw1;       !构建调节项
         xwyw2 | xw2 XWITH yw2;
```

```
        xwyw3 | xw3 XWITH yw3;
        xw2 on xwyw1(cx);
        xw3 on xwyw2(cx);
        xw4 on xwyw3(cx);
        yw2 on xwyw1(cy);
        yw3 on xwyw2(cy);
        yw4 on xwyw3(cy);
OUTPUT: TECH8;
PLOT:TYPE = PLOT2;
```

注：如果要进行跨层次调节效应分析(以 $X_{Bi}Y_{Wi(t-1)}$为例)，只需要添加如下语句

```
    xbyw1 | xb XWITH yw1; !构建调节项
    xw2 on xbyw1(c1x);
```

基于多层结构方程模型的调节效应分析的 Mplus 程序：

```
DATA:          FILE = 1.csv;
VARIABLE:      NAMES = x z y person;
               CLUSTER = person;
ANALYSIS: TYPE = TWOLEVEL RANDOM;
               ESTIMATOR = BAYES;
               ALGORITHM = GIBBS(RW);
               POINT = MEAN;   !用均值作为点估计
               CHAINS = 2;
               PROCESSORS = 3;
               BITERATIONS = 50000 (20000);   !马尔科夫链最少迭代 2 万次，最多 5 万次
               THIN = 10;
MODEL:   %within%
               xw by x@1; x@.01;
               zw by z@1; z@.01;
               s1 | y on xw;
               s2 | y on zw;
               xw with zw;
               xzw | xw XWITH zw;   !构建个体内的调节项
               y on xw zw xzw;   ! LMS 法求调节效应 γ30
         %between%
               xb by x@1; x@.01;
               zb by z@1; z@.01;
               s1 on zb;   !随机系数预测法求调节效应 γ11
               xzb | xb XWITH zb;   !构建个体间的调节项
               y on xb zb xzb;   !LMS 法求调节效应 γ03
               xb with zb;
OUTPUT:   CINTERVAL(hpd) TECH8;
plot: type = plot2;
```

基于潜变量增长模型的调节效应分析的 Mplus 程序：

```
DATA:           FILE = 1.dat;
VARIABLE:       NAMES = x1-x4 z1-z4 y1-y4;
ANALYSIS:       ESTIMATOR = BAYES;
                TYPE = RANDOM;
                POINT = MEAN;
                CHAINS = 3;
                PROCESSORS = 3;
                BITERATIONS = (20000);
                BCONVERGENCE = 0.025;   !马尔科夫链收敛标准
                THIN = 10;
MODEL:          IX SX | x1@0 x2@1 x3@2 x4@3;        !构建 LGM 模型
                IZ SZ | z1@0 z2@1 z3@2 z4@3;
                IY SY | y1@0 y2@1 y3@2 y4@3;
                int1 | SX XWITH SZ;     !构建变化率的调节项
                int2 | IX XWITH IZ;      !构建初始水平的调节项
                SY ON SX SZ int1;
                IY ON IX IZ int2;
OUTPUT:     CINTERVAL(hpd)    TECH8;    STDYX;
PLOT: TYPE = PLOT2;
```